PCI-SIG ENGINEERING CHANGE REQUEST

TITLE:	Enabling Multiple Base Addresses per PCI Segment Group
DATE:	August 13, 2018
AFFECTED DOCUMENT:	PCI Firmware Specification, Rev. 3.2
SPONSOR:	Jim Panian, Qualcomm Technologies, Inc. Joe Cowan, Hewlett Packard Enterprise

Part I

1. Summary of the Functional Changes

Changes are made to Sections 4.1.2 through 4.1.4, impacting the MCFG Table Description and _CBA Method.

The changes enable the MCFG table format to allow for multiple memory mapped base address entries, instances of Table 4-3, per single PCI Segment Group.

Currently, the MCFG table format allows for more than one memory mapped base address entry, instance of Table 4-3, provided each entry corresponds to a unique PCI Segment Group consisting of 256 PCI buses. Multiple entries corresponding to a single PCI Segment Group are not allowed.

Changes to the _CBA Method enable the same model, namely that different bus number ranges within a single PCI Segment Group may have a different configuration base address associated with each bus number range.

2. Benefits as a Result of the Changes

Modern operating systems enable the 256 MB of PCIe Configuration Space per PCI Segment Group to be memory mapped in a non-contiguous fashion through the technique proposed by this ECR. Currently, the PCI Firmware Specification is written assuming that the 256 MB of PCIe Configuration Space per PCIe Segment Group is memory mapped in a contiguous fashion.

3. Assessment of the Impact

ACPI must be enhanced to interpret multiple instances of “Table 4-3: Memory Mapped Enhanced Configuration Space Base Address Allocation Structure” per PCI Segment Group.

4. Analysis of the Hardware Implications

No impact.

5. Analysis of the Software Implications

The MCFG table is an ACPI table that is used to communicate the base addresses corresponding to the non-hot removable PCI Segment Groups range within a PCI Segment Group available to the operating system at boot. The operating system must be able to interpret the changes.

6. Analysis of the C&I Test Implications

No impact.

Part II

Detailed Description of the change

Change Section 4.1.2 page 42 as follows:

4.1.2. MCFG Table Description

The MCFG table is an ACPI table that is used to communicate the base addresses corresponding to the non-hot removable PCI Segment Groups range within a PCI Segment Group available to the operating system at boot. This is required for the PC-compatible systems.

The MCFG table is only used to communicate the base addresses corresponding to the PCI Segment Groups available to the system at boot. This table directly refers to PCI Segment Groups defined in the system via the _SEG object in the ACPI name space for the applicable host bridge device. For systems containing only a single PCI Segment Group, the default PCI Segment Group number, namely, PCI Segment Group 0, is implied. In such a case, the default PCI Segment Group need not be represented in the ACPI Name Space (i.e., no _SEG method is required in such a hierarchy).

The size of the memory mapped configuration region is indicated by the start and end bus number fields in the Memory mapped Enhanced configuration space base address allocation structure as shown in Table 4-3. 0-255 is the range of allowed bus numbers supported for a given PCI Segment Group.

Table 4-2 provides a description of the MCFG table.

Table 4-2: MCFG Table to Support Enhanced Configuration Space Access

Field	Byte Length	Byte Offset	Description
Header			
Signature	4	0	"MCFG". Signature for the Memory mapped configuration space base address Description Table. (refer to Note 1)
Length	4	4	Length, in bytes, of the entire MCFG Description table including the memory mapped configuration space base address allocation structures.
Revision	1	8	1
Checksum	1	9	Entire table must sum to zero
OEMID	6	10	OEM ID
OEM Table ID	8	16	For the MCFG Description Table, the table ID is the manufacture model ID
OEM Revision	4	24	OEM revision of MCFG table for supplied OEM Table ID
Creator ID	4	28	Vendor ID of utility that created the table

Creator Revision	4	32	Revision of utility that created the table
Reserved	8	36	Reserved
Configuration space base address allocation structure [n]	---	44	A list of the memory mapped configuration base address allocation structures. This list will contain at least one entry corresponding to each PCI Segment Group present in the platform. The structure of this entry is defined in Table 4-3.

Notes:

1. A table signature "MCFG" is reserved for this purpose and the header for the table is shown in Table 4-2. Based on the signature and table revision, the operating system can then interpret the implementation- specific data within the table. The Table Revision for revision 1.0 of the MCFG table is set to 1.
2. If the operating system does not natively comprehend reserving the MMCFG region, the MMCFG region must be reserved by firmware. The address range reported in the MCFG table or by _CBA method (see Section 4.1.3) must be reserved by declaring a motherboard resource. For most systems, the motherboard resource would appear at the root of the ACPI namespace (under _SB) in a node with a _HID of EISAID (PNP0C02), and the resources in this case should not be claimed in the root PCI bus's _CRS. The resources can optionally be returned in Int15 E820h or EFIGetMemoryMap as reserved memory but must always be reported through ACPI as a motherboard resource.
3. This table must not include the memory mapped configuration base addresses for hot pluggable PCI Segment Groups. Such PCI Segment Groups must be described by using the _CBA method (see Section 4.1.3) in the corresponding ACPI name space object.

The structure in Table 4-3 describes the association between the PCI Segment Group and the corresponding memory mapped configuration base address. This table describes the details of this structure.

Table 4-3: Memory Mapped Enhanced Configuration Space Base Address Allocation Structure

Field	Byte Length	Byte Offset	Description
Base Address	8	0	Processor-relative Base Address for the Enhanced Configuration Access Mechanism
PCI Segment Group Number	2	8	PCI Segment Group Number. Default is 0. For all other PCI Segment Groups, this field value should correspond to the value returned by _SEG object in ACPI name space for the applicable host bridge device.
Start Bus Number	1	10	Start PCI Bus number decoded by the host bridge
End Bus Number	1	11	End PCI Bus number decoded by the host bridge
Reserved	4	12	Reserved

The MCFG table format allows for more than one memory mapped base address entry (instance of Table 4-3) provided each entry (memory mapped configuration space base address allocation structure) corresponds to a unique PCI Segment Group consisting of 256 PCI buses. Multiple entries corresponding to a single PCI Segment Group ~~is not~~ are also allowed provided <PCI

Segment Group Number, Start Bus Number, End Bus Number> uniquely identifies each PCI Host Bridge and the bus number values do not overlap.

Change Section 4.1.3 page 43 as follows:

4.1.3. The _CBA Method

Some systems may support hot plug of host bridges that introduce either a range of buses within an existing PCI Segment Group or introduce a new PCI Segment Group. For example, each I/O chip in a multi-chip PCI Express root complex implementation could start a new PCI Segment Group. The base address of the memory mapped configuration space for such a hot pluggable PCI Segment Group or a range of buses within a PCI Segment Group is described using an ACPI control method, _CBA, that is under the host bridge devices that are part of the PCI Segment Group. This applies to PC-compatible systems only.

The _CBA (Memory mapped Configuration Base Address) control method is an optional ACPI object that returns the 64-bit memory mapped configuration base address for the hot plug capable host bridge. The base address returned by _CBA is processor-relative address. The _CBA control method evaluates to an Integer.

This control method appears under a host bridge object. When the _CBA method appears under an active host bridge object, the operating system evaluates this structure to identify the memory mapped configuration base address corresponding to ~~the PCI Segment Group for~~ the bus number range specified in _CRS method. Within the same PCI Segment Group, different host bridges, each with its associated bus number range, may have a different configuration base address. An ACPI name space object that contains the _CBA method must also contain a corresponding _SEG method.

For a host bridge that includes _CBA, the _CBA and _BBN control methods have to be executed first to enable PCI_Config_OpRegion access for devices below the bridge. As a result, the _CBA and BBN methods must not include PCI_Config opregions that refer to devices below the host bridge.

A set of hot pluggable host bridges could have _CBA under each of the host bridge devices, where each host bridge device is typically described in the ACPI name space with PNP0A08 for _HID and PNP0A03 for _CID. In this case, the memory mapped configuration base address (always corresponds to bus number 0 of a given bus number range) ~~for the PCI Segment Group of the host bridge~~ is provided by _CBA and the bus number range covered by the base address is indicated by the corresponding bus number range specified in _CRS.

If rebalancing of resources on a host bridge is supported via _PRS, _SRS, it is the responsibility of the operating system to reevaluate _CBA every time _CRS is evaluated.

Change Section 4.1.4 page 44 as follows:

4.1.4. System Software Implication of MCFG and _CBA

The base address returned by MCFG table for a given ~~PCI Segment group~~ bus number range is always with respect to bus 0 of that particular bus number range as specified in the PCI Express Base Specification (and the PCI-X Specification). It is the responsibility of system software to calculate the start and end of the supported memory mapped configuration address range based on the start and end bus numbers specified in the MCFG entry. System software must make no

assumptions about the memory range corresponding to the base address up to the start of the memory mapped configuration space (as specified by start bus number).

…

IMPLEMENTATION NOTE

Multiple Host Bridges

A platform may have multiple PCI Express or PCI-X host bridges. The base address for the MMCONFIG space for these host bridges may need to be allocated at different locations. Historically, ~~In~~ in such cases, using MCFG table and _CBA method as defined in this section means that each of these host bridges must be in its own PCI Segment Group. This approach is referred to as the legacy interpretation of this specification. A newer interpretation of this specification allows "multiple host bridges with base addresses allocated at different locations" to exist within the same PCI Segment Group. Vendors who choose to implement to the newer interpretation bear the responsibility that supported operating systems can handle the newer interpretation.

Allocating base addresses to non-contiguous regions of memory has been achievable in some implementations with an address mapping feature to make non-contiguous regions of memory appear to be contiguous to the operating system. The newer interpretation removes the need for such an address mapping feature thereby simplifying implementations.

Allocating base addresses to non-contiguous regions of memory has also been possible using the legacy interpretation of the specification by employing different PCI Segment Group Numbers. However, the newer interpretation of the specification preserves the use of a single PCI Segment Group Number since implementations will not allow peer-to-peer communications across multiple PCI Segment Groups.

For example, the Base Address field value returned by MCFG of a bus number range starting at 0xE2000000, where the Start Bus Number is 0, would be 0xE2000000. The Base Address field value returned by MCFG of a bus number range starting at 0xE2000000, where the Start Bus Number is 0x40, would be (0xE20000000 – (0x40 << 20)) or 0xDE000000. Therefore, the lowest possible Base Address field value returned by MCFG is a function of Start Bus Number:

Base Address = (Start of memory mapped configuration address range)-((0x100000)*(Start Bus Number)).

If this platform also needs to support legacy operating systems or x86 BIOS Option ROMs, the CF8/CFCh access mechanism (which is PCI Segment Group unaware and only can support up to 256 bus numbers) must also be supported. There may be a number of implementation choices to make these two work together; for example, the bus numbers in each of the PCI Segment Group can be made to not overlap. This would make the CF8/CFCh access still appear to be Segment Group unaware and support up to 256 buses while using the MCFG/_CBA definition to describe host bridges MMCONFIG base addresses that are allocated at different locations.

Note that in this arrangement, even though the PCI Segment Group concept is used, the total number of PCI buses is still limited to 256 due to the CF8/CFCh limitation.
